DOCUMENT RESUME

ED 403 492 CE 073 511

TITLE	Evaluation Systems in HRD.
PUB DATE	96
NOTE	25p.; In: Academy of Human Resource Development (AHRD) Conference Proceedings (Minneapolis, MN, February 29-March 3, 1996); see CE 073 480.
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Research/Technical (143) -- Reports - Descriptive (141)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Adult Education; Case Studies; *Evaluation Methods; Institutional Evaluation; *Labor Force Development; Models; *Performance Technology; *Personnel Evaluation; Pilot Projects; *Program Evaluation; *Systems Approach
IDENTIFIERS	External Evaluation; Internal Evaluation; Strategic Knowledge

ABSTRACT

This document contains three papers presented at a symposium on evaluation systems in human resource development (HRD) moderated by Jerry Gilley at the 1996 conference of the Academy of Human Resource Development. "PLS [Performance-Learning-Satisfaction] Evaluation System: Sales Communication Case Study" (Richard A. Swanson, Timothy R. McClernon) is a case study that evaluates the workability of the three-domain PLS evaluation model by applying it to a performance improvement intervention focused on sales communication training. John C. Redding's paper, "Assessing Organization, Team, and Individual Learning: The Strategic Learning Assessment Process" describes the development and pilot application of an organizational learning assessment process based on the Strategic Learning Model that Redding and Catalanello proposed in 1994. "The Realities of Communicating and Reporting Practices and Outcomes for Internal and External Evaluators" (Rosalie T. Torres, Hallie Preskill, Mary E. Piontek) reports the results of a study in which 240 of a random sample of 343 American Evaluation Association members responded to a survey designed to identify internal and external evaluators' practices and concerns about communicating and reporting evaluation findings. Papers contain references. (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 403 492

SYMPOSIUM
[30] EVALUATION SYSTEMS IN HRD

Chair: Jerry Gilley, Western Michigan University

PLS Evaluation System: Sales Communication Case Study
Richard A. Swanson, University of Minnesota
Timothy R. McClernon, CIGNA

Assessing Organization, Team, and Individual Learning: The Strategic Learning Assessment Process
John C. Redding, Northern Illinois University

Realities of Communicating & Reporting Practices and Outcomes for Internal and External Evaluators
Rosalie T. Torres, Torres Consulting Group
Hallie Preskill, University of New Mexico
Mary E. Piontek. RMC Research Corp.

Academy of Human Resource Development
1996 Conference
Minneapolis, MN
February 29-March 3, 1996

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

R. Swanson, J. Redding, R. Torres

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

CE073511

BEST COPY AVAILABLE

PLS Evaluation System: Sales Communication Case Study

Richard A. Swanson
University of Minnesota

Timothy R. McClernon
CIGNA

This case study is an application of the three-domain performance-learning-satisfaction (PLS) evaluation model to a performance improvement intervention focused on sales communication training. The case study inquiry questions revolved around workability of the PLS, in-process evaluation activities and decisions, and the actual evaluation result. Among the results, the Sales Communication training yielded a net financial performance gain of $2,352,300 and a 8:1 financial return on investment in less than a year.

The gap between sound evaluation theory and typical practice is a serious problem for human resource development-performance improvement (HRD-PI) in industry, business, and government organizations (Swanson, 1990). HRD-PI includes personnel training and development, organization development, and more general performance improvement efforts. Evaluation is one of the primary phases of the systematic HRD-PI process. The first four phases of the process include analysis, design, development and implementation (or some variation of these elements).

A literature review of the summative evaluation noted that evaluation, when conducted, is not conducted effectively (Parker, 1986). A typical work-place scenario consists of the busy HRD-PI practitioner doing what the company wants, feeling successful, and not being regularly required to prove the added value that results from HRD-PI contributions. With a full agenda of important development and delivery tasks, the busy professional finds it difficult to evaluate. However, almost all important organizational processes and functions are regularly evaluated in terms of their effectiveness, efficiency, and bottom-line contributions to the enterprise. In addition, it has been clearly established that effectiveness evaluation data, particularly bottom-line performance results, are the key to gaining support for the HRD-PI function from top management (Kusy, 1988). It is clearly irrational *not* to evaluate HRD-PI programs and processes.

Both Parker's review of literature and Kusy's study of management support of effectiveness evaluation established the need for this study. The long-standing and popular four-level evaluation model that dominates the profession has been judged by most thoughtful minds be invalid and invalid, causing problems rather than solving problems for the profession (Alliger & Janak, 1989; Holton, 1996; Newstrom, 1995). Sadly, the most vocal defenders of inadequate theory and practice demonstrate little understanding of sound evaluation practice or theory while advocating for simplicity (Kirkpatrick, 1995 & 1996). It is important to note that and evaluation can be both simple and valid, and that a simple and invalid evaluation effort is worse than no formal evaluation

Purpose of the Study

The purpose of this manuscript is to report one of a series of case studies applying the Performance-Learning-Satisfaction Evaluation System (PLS Evaluation System) (Swanson, 1995). The inquiry questions include: (1) Can the PLS Evaluation System be applied to the sales communication case? (2) What in-process activities and/or decisions in the application of the PLS Evaluation System to the sales communication case provide additional understanding? (2) What are the PLS Evaluation System data resulting from the sales communication case study?

© *copyright, Swanson & McLernon, 1996*

PLS System Overview

The *PLS Evaluation System* addresses the evaluation phase HRD-PI and consists of five major elements: (1) PLS Evaluation *Model*, (2) PLS Evaluation *Plan*, (3) PLS Evaluation *Tools*, (4) PLS Evaluation *Schedule*, and (5) PLS Evaluation *Report*.

The PLS Evaluation System is grounded in systems, psychological, and economic theories. The evaluation system logically connects the performance goals specified in the up-front performance analysis with the performance outcomes. Also acknowledged are the related or mediating outcomes of learning and satisfaction (Swanson, 1995).

The System contends that without a connection to up-front analysis, it is unlikely that performance results at the individual, process, or organizational levels will be found. The fundamental flaw of providing an intervention and then looking for results (or transfer) has been at the crux of HRD-PI poor practice (Swanson, 1994). Even so, the results of placing a rigorous evaluation process on top of a poorly conceived HRD-PI intervention can minimally cause an organization to reassess its practices.

The *PLS Evaluation System* purports to be a rigorous, yet flexible, evaluation system. It is flexible enough to accommodate almost all situations. It is rigorous enough to consistently honor core questions, techniques, and reporting results as part of every evaluation. In addition to the *PLS Evaluation System* conceptual components, job aids, tools, and references, there is an auxiliary data processing system consisting of a computer, software, a data input scanner, and a printer. This system allows for the production of program specific scannable data entry sheets, scanning of completed sheets into the computer memory, and the manipulation of the data into a standard format report.

PLS Evaluation Model

The evaluation model embraces the domains of performance, learning, and satisfaction (PLS). Each of the three domains are divided into two options for a total of six evaluation options labeled A-F.

The three domains of performance, learning, and satisfaction are not levels or a hierarchy (see Holton, 1996; Newstrom, 1995). They are domains of independent worth from an evaluation perspective. While there can be a relationship between the domains, it cannot and should not be assumed that there is a direct positive relationship. Contrary to the practitioners myth, the research shows that participants most satisfied with a program are not necessarily those who learned the most (Alliger & Janak, 1989; Dixon, 1990). High or low satisfaction can be found among low, medium, and high achievers. And, because participants have gained knowledge and expertise does not mean that they will use it in the workplace (Gielen, 1995).

Performance can be thought of as (a) business results at the organizational, process or individual levels and/or (b) financial results or benefits in terms of money or monetary ratios. Most, but not all, business results can be monatized and turned into financial results. For example, the business result of increasing market share by 5% could be converted into numbers of sales times the financial value of each sale. Another example would be to increase the hourly production rate of injection molded items (that meet quality standard) from 30 to 35. In each of these examples the business results can be monitized and then expressed in terms of financial results.

The general financial results goal of the *PLS Evaluation System* is that the benefits exceed the costs by a 2:1 ratio. If the 2:1 return on investment goal is achieved, it can be said that the program achieved 100% of its goal. A 2:1 return on investment, especially in a year or less, is ambitious. Organizations may want to choose and adhere to a lesser goal. In all instances the actual data can be reported as well as the goal attainment data.

Learning can be thought of as (c) knowledge demonstrated through completion of tests of information and concepts and/or (d) expertise in demonstrating the skillful workplace behaviors in a simulated or actual work situation. Knowledge is most efficiently measured through objective tests presented through paper and pencil or technology. This is because objective tests allow for a larger sample of learning from each participant in a minimum amount of time.

While knowledge is a component of expertise, demonstration and assessment of skillful workplace behaviors pushes the training closer to reality. As with the performance and satisfaction

domains, rational goals need to be established for the learning areas of knowledge and expertise. The 100% goals for knowledge and/or expertise must be connected to the behavioral outcomes established for the intervention. These outcomes generally focus on (1) not meeting, (2) meeting, or (3) exceeding standards. The general learning results goal of the *PLS Evaluation System* is that participants, on the average, meet learning and expertise standards. Thus, if meeting standard is 2 on a 1-3 scale, an average score of 2 would be attainment of 100% of the goal.

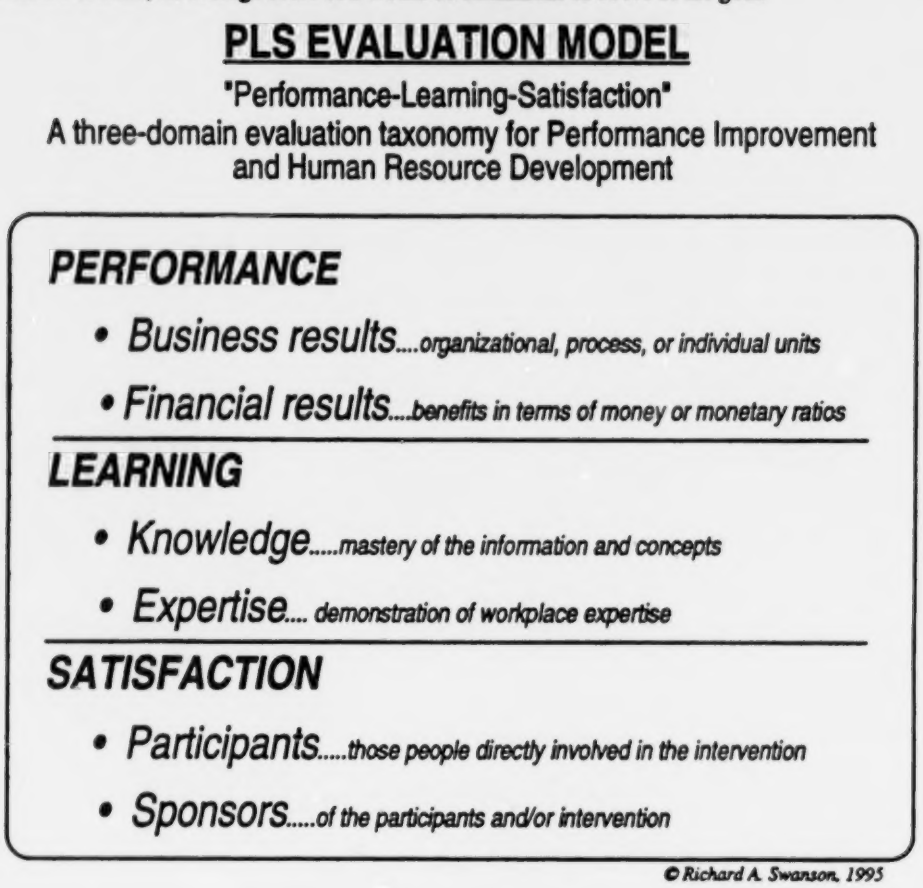

PLS EVALUATION MODEL
"Performance-Learning-Satisfaction"
A three-domain evaluation taxonomy for Performance Improvement and Human Resource Development
PERFORMANCE
• Business results....organizational, process, or individual units
• Financial results....benefits in terms of money or monetary ratios
LEARNING
• Knowledge.....mastery of the information and concepts
• Expertise.... demonstration of workplace expertise
SATISFACTION
• Participants.....those people directly involved in the intervention
• Sponsors.....of the participants and/or intervention
© Richard A. Swanson, 1995

Figure 1. PLS Evaluation Model

Satisfaction can be thought of as the (e) perceptions of the interventions by the participants directly involved and/or (f) perceptions of the sponsor (or supervisor) as to the participants and/or interventions. The overriding concern is that perceptions be acceptable, not negative and not jubilant. Since most development and improvement efforts require change, and change is not comfortable, it is illogical to pursue extremely high ratings. And, extremely low ratings end up being culturally counter productive. The standard 4 Likert scale (2 levels each of negative and positive with an acceptable mid-point score of 2.5) establish 2.5 average as the "acceptable" satisfaction goal. Thus, if the satisfaction rating average is 2.5, it can be said that the program achieved 100% of its goal. Only after there is clear evidence that performance and learning goals are being reached should there be focused concern for raising the 2.5 satisfaction goal.

Case Study Context

The application of the *PLS Evaluation System* in this case study involved a communication training program for sales personnel in a major insurance company. The salesforce sells insurance

plans to employers. The performance improvement intervention in this case included training that was provided to 167 sales personnel through 24 program offerings at various sites throughout the nation over a six month period.

The company had an ongoing commitment to generic communication training for their sales force. Given this base of support, the manager conducted a performance diagnosis that confirmed the loss of business as a result of inadequate communication skills and a more precise profile of the specific communication expertise required. Supervisors of the sales personnel were also trained in coaching the communication skills, were recruited to assist in the role-playing assessments of expertise during training, and in conducting follow-up evaluations and coaching back on the job.

The company has a system for tracking sales by salesperson, office, region, and product. They also have standard gross and net financial measures per sale. The connection between available financial measures and the specific of this case were not known at the onset. They became clear through the application of the *PLS Evaluation System*.

The inquiry questions directly the case study were: (1) Can the PLS Evaluation System be applied to the sales communication case? (2) What in-process activities and/or decisions in the application of the PLS Evaluation System to the sales communication case provide additional understanding? (2) What are the PLS Evaluation System data resulting from the sales communication case study?

Case Study Evaluation Plan Results

In the *PLS Evaluation System*, planning decisions are recorded on the *PLS Evaluation Plan*. Figure 2 is a sample completed *PLS Evaluation Plan* for "Communication Training" .

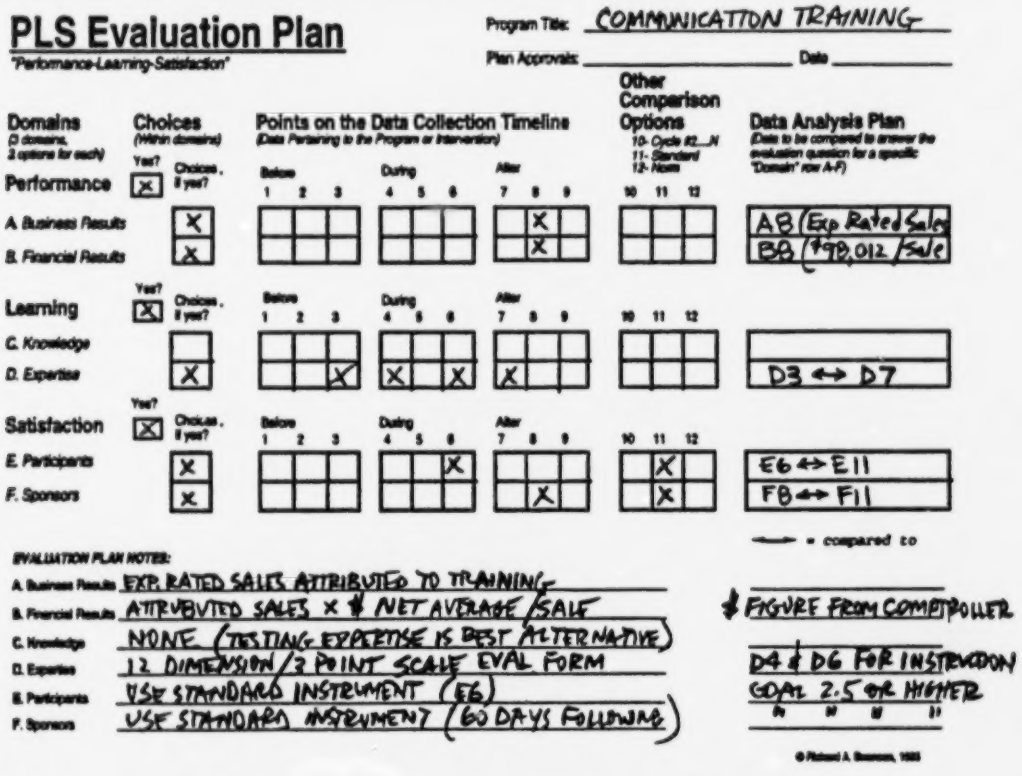


Figure 2. PLS Evaluation Plan for the Sales Communication Intervention

BEST COPY AVAILABLE

The "Domains & Options" rows of the plan follow the three evaluation domains and the six evaluation options presented in the *PLS Evaluation Model* (Figure 1). The vertical columns require the planner to make a series of decisions in relation to the evaluation domains and options. Once the evaluation option choices are made, data pertaining to the program or intervention is collected within a timeline. This timeline can generally be thought of as being before, during, or after the intervention.

All three domains were checked for the Communication Training case. Of the six A-F options, only "C. Knowledge Test" was not checked. This was not checked because of the very tight connection of the content of the training program to the test of expertise, option D, and because of the practical concern for the amount of in-program time that could be dedicated to testing.

For "A. Business Results," the decision was to measure sales that were attributed to individuals and to do this 60 days following the training. This was the most complex decision in the plan. In this case setting sales personnel make a small number of sales that are individually worth a great deal. Time needed to lapse in order to determine the effect, if any, the communication training had on sales. Unlike other case studies, there was not a reasonable comparison group given the changing market, limited sales, and across-the-board training that took place. Thus, it was determined that sales people and their supervisors would be asked to report on any sales, or portion of sales, attributable to the training. The position taken was these estimates would be underestimates and that the specific sale had to be recorded if there were any questions about the data base. Thus, these were not perceptions of performance. Rather, they were specific actual sales.

For "B. Financial Results" the total number of sales was multiplied by the net worth of an average sale, $120,000. A meeting with the comptroller yielded this figure following a zigzag conversation about available financial data. The direct cost of the training is compared to the net profit from the reported sales attributable to Communication Training to determine the ROI.

For "D. Expertise" an instrument was created to assess expertise. This simple rating instrument was used before, during, and after the Communication Training and was instrumental in connecting the training and line management elements of this performance improvement intervention. Trained observers rated participants in a role playing situation used a 10 category instrument with each category assessed as exceeds standards (3 points, meets standards (2 points), and below standards (1 point). The goal was to obtain an average score of 2.0 at the end of training, D6 on the timeline. In this case the average score was 1.89 or 98% of the goal. Also reported were the before (D3) and after (D7) scores which were 78% and 103% of goal.

For "E. Participant Satisfaction" a standardized PLS rating form was completed by each participant at the completion of training. These were averaged and compared to the standard of 2.5 or higher on a 1-4 scale.

For "F. Sponsor Satisfaction" a standardized PLS rating form was completed by each supervisor of participants sixty days following the completion of training (F8). These were averaged and compared to the standard of 2.5 or higher on a 1-4 scale (F11).

Case Study Evaluation Tools

Evaluation tools can be constructed and used for each of the six evaluation options (A-F).
Of these, performance scores are indicators of the business effects that result from an HRD program. Learning scores are indicators of the amount of knowledge and expertise acquired by the participants during the development program, and the satisfaction scores are indicators of how participants and their sponsors perceive the program. Although there are many possible ways to configure evaluation tools, *the PLS Evaluation System* focuses on six reasonable options (A-F) and recognizes that not all six are always used.

Performance. In the *PLS Evaluation System*, the tools for measuring the performance results back in the organization are business results and financial results. Business results contrast the productivity of either the organization, business process, or individual employees (1) before and after the HRD-PI program or (2) against a comparison group, standard, or norm. It is always desirable to utilize the organization's existing direct measures of performance such as sales, items produced, customer's served. In that organizations are dynamic, judgments may be required in terms of attributing the portion of performance attributable to the intervention being evaluated. In

BEST COPY AVAILABLE

this case, and in every case I have participated in, estimates of the HRD-PI contribution by line managers have been underestimates. As might be expected, line personnel tend to want to take as much credit as they can. Even so, this state of affairs still leaves high quality HRD-PI interventions in the realm of excellent business investments.

Financial results are used to determine the economic value of the HRD program, the benefit of which is determined by subtracting the cost of the program from the performance value resulting from the program. The financial cost accounting methods and norms utilized by the organization should be consistently applied to all calculations. Thus, if employee salary during the time they are participating in a development program is not normally costed back to the intervention, it is not considered as part of the cost basis. In many calculations this salary cost variable will already show itself in lost productivity from being absent from the workplace.

On the performance value side, it is generally easier than you think to determine the value of a unit of performance. This critical financial information is almost always a few tiers away from the place where the actual performance takes place. Financial benefits are a matter of subtracting costs from the resulting performance values. Interpretation of this data has been discussed in earlier sections of this article.

Learning. Learning is measured by knowledge tests, tests of expertise, or both. Knowledge tests measure the cognitive information learned by participants. Two types knowledge test items, multiple choice and matching, are encouraged because they can be scored objectively and are not as susceptible to guessing as are other forms of test items and not as subjective and time consuming as verbal testing or written essays. In constructing knowledge tests, as with all tests, care must be taken to ensure that the test produces valid and reliable results. A test is valid when it measures what it is supposed to measure and it is reliable when it produces consistent results.

Tests of expertise measure what the learners can do by examining either the products that the learners produce or the processes exhibited by the participants as a result of their learning. An in-program test of expertise must also be valid and reliable.

Satisfaction. The *PLS Evaluation System* requires that "satisfaction" be measured for every intervention. Participant satisfaction is measured by having each participant complete the Program Evaluation Form at the completion of the program. This is a standardized form that should not be altered in terms of the front page questions. The participant satisfaction score is calculated by tallying all the participant responses to the front page questions. Ordinal values are then assigned to descriptors as follows: very good (4), good (3), fair (2) and poor (1). The overall participant satisfaction score is obtained by averaging the scores and determining the mean satisfaction score which will fall within the 1-4 range. Sub-scores on the individual questions can also be computed this way.

Comments written by the participants on the participant satisfaction form are not included in the participant satisfaction score. They provide open-ended feedback for the facilitators, presenters, and program planners for program improvement and are used to highlight the report.

Sponsor satisfaction is measured by using the Sponsor Evaluation Form (Figure 8), which is completed by each participant's supervisor or sponsor. After the responses are gathered, the average sponsor satisfaction score for the HRD-PI program is computed in the same manner that the average participant satisfaction score and sub-scores are determined. As before, the written comments provide the facilitator, presenter, or program planner with immediate, open-ended feedback.

Using standard participant and sponsor satisfaction forms for all training courses allows for the comparisons of HRD-PI programs with each other and across time. This makes it easy to identify and document recurring problems and/or successes. The use of scannable participant data sheets makes the data analysis quite easy.

Case Study Evaluation Report

The *PLS Evaluation Report* is an "Executive Summary" of the effectiveness of a program (or set of identical interventions). The intention is that every program offering be evaluated in terms of its' effectiveness and that the results be reported to the appropriate stakeholders in the organization. *PLS Evaluation Reports* have a standard format, standard sections, and standard means of reporting data. And, reports are short, generally 2-4 pages in length.

There are eight standard sections: (1) Organization and Program Identification Heading, (2) Program Purpose, (3) Program Description, (4) Evaluation Summary, (5) Approval, (6) Distribution List, (7) Evaluation Results (Performance, Learning, Satisfaction, and (8) Improvement Proposal.

In that the *PLS Evaluation Report* is an executive summary, there is available evaluation data exceeding what is contained in the report. This additional data is retained and used by the department for tracking and improving specific elements of the program and for responding to specific evaluation inquiries. In summary, the evaluation results for Communication Training were as follows:

A. Performance, Business Results: In total, 24 sales by the participants in the 24 training sessions were verified as having been the direct result of the communication training in the nine months following the training. No comparison to goals or norms was made in that the annualized data is still unavailable.

B. Performance, Financial Results: The net profit from the reported sales attributable to Communication Training was $2,352,3000. The cost of the 24 training sessions was $288,000. The conservative estimate of the financial return-on-investment in the six months was 8:1 ROI. The standard PLS Evaluation System financial goal is 2:1. Thus, the Communication Training attained 409% of goal for the nine month period.

C. Learning, Knowledge: No knowledge tests were administered as it was determined that the tests of expertise directly verified the required knowledge and that having one in-training test was adequate.

No comparison made

D. Learning, Expertise: Using a standardized rating form, expert raters evaluated the expertise of the participants on a multi-dimensional instrument. A mean rating of 2.0 was equivalent to "meets clients expectations." Before training the average participant rating was 1.71 (86% of goal), at the end of training it was 2.12 (106% of goal), and 60 days after training it was 2.33 (117% of goal).

E. Satisfaction, Participant: The average participant satisfaction rating by 167 people on a 4 point scale at the end of program on was 3.72 compared to the goal of 2.5. Thus Communication Training attained 149% of the participant satisfaction goal.

F. Satisfaction, Sponsor: The average sponsor satisfaction rating 60 days after the program was 3.68 compared to the goal of 2.5. Thus, Communication Training achieved 147% of sponsor satisfaction goal.

The PLS Evaluation System is designed in a manner that allows the utilization of computer hardware and software to create instruments, scan data into the computer, and to call out the data required for the PLS Evaluation Report. These were all utilized in the case study and still require further development.

Conclusion

The purpose of this manuscript was to present a case study of the application of a purported practical and valid evaluation system that can be applied to any HRD-PI program in business, industry, and government-- for profit or non-profit organizations.

In this case study the PLS Evaluation System was successfully applied to the sales communication training and the data provided convincing evidence of the effectiveness of the sales communication intervention.

While the case study validates the workability of the PPS Evaluation System, several issues emerged and are presented in the form of questions for further research and/or development:

- Should the person doing the up-front performance diagnosis also be responsible for the performance evaluation component of the PLS evaluation?
- Should the primary internal customer of the evaluation report be involved in reviewing the evaluation plan?
- What minimum level of knowledge and expertise should a person have before managing any individual evaluation effort?
- What are the response from various stakeholders to the practice of setting a 100% goal for each evaluation area?
- What is the organization response to PLS evaluations over time?

BEST COPY AVAILABLE

References

Alliger, G. M. & Janak, E. A. (1989). Kirkpatrick's levels of training criteria: Thirty years later. Personnel Psychology, 42, 331-340.

Dixon, N.M. (1990). Relationship between trainees responses on participant reaction forms and their posttest achievement score. Human Resource Development Quarterly, 1 (2).

Holton, E. F. (1996). The flawed four-level evaluation model. Human Resource Development Quarterly, 7 (1).

Kearsley, G. (1982). Costs, benefits, and productivity in training systems. Reading: Addison-Wesley.

Kusy, M.E. (1988). The effects of types of training on support of training among corporate managers. Performance Improvement Quarterly, 1 (2), 23-30.

Kirkpatrick, D. L. (1996). Reaction to Holton. Human Resource Development Quarterly. 7 (1).

Mosier, N.R. (1990). Financial analysis: The methods and their application to employee training. Human Resource Development Quarterly, 1 (1), 45-63.

Parker, B.L. (1986). Summative evaluation in training and development. Journal of Industrial Teacher Education, 23 (2), 29-55.

Robinson, D. & Robinson, J. (1989). Training for impact: How to link training to business needs and measure the results. San Francisco: Jossey-Bass

Swanson, R.A. (1989). Everything important in business is evaluated. In R.O. Brinkerhoff (Ed.) New Directions in Program Evaluation - Evaluating Training Programs in Business and Industry. San Francisco; Jossey-Bass, 44, 71-82.

Swanson, R. A. (1992). Demonstrating financial benefit to the client. In Stolovitch, H.D. & Keeps, E.J. (Eds.). Handbook of human performance technology. San Francisco: Jossey-Bass.

Swanson, R. A. (1994). Analysis for improving performance: Tools for diagnosing organizations and documenting workplace expertise. San Francisco: Berrett-Koehler.

Swanson, R. A. (1995). Human resource development: Performance is the key. Human Resource Development Quarterly. 6 (2), 207-213

Swanson, R. A. (1995). Performance-learning-satisfaction evaluation system: Three-domain model for evaluating performance improvement, human resource develolment and quality improvement interventions. University of Minnesota, St. Paul: Human esource Development Research Center.

Swanson, R.A. & Gradous, D.B (1988). Forecasting financial benefits of human resource development. San Francisco, CA: Jossey-Bass.

Richard A. Swanson is Professor and Director of the Human Resource Development Research Center, University of Minnesota, St. Paul, MN.

Timothy R. McClernon is Vice President of Sales Performance Consulting, CIGNA, Bloomfield, CT

Assessing Organization, Team, and Individual Learning: The Strategic Learning Assessment Process

John C. Redding
Institute for Strategic Learning
Northern Illinois University

This paper describes the development and pilot application of an organizational learning assessment process based upon the Strategic Learning Model proposed by Redding and Catalanello (1994). It presents an overview and case illustration of the assessment process. It then describes the first phase of a three-phase pilot study. Findings support the practicality of the approach, while suggesting improvements to increase assessment reliability.

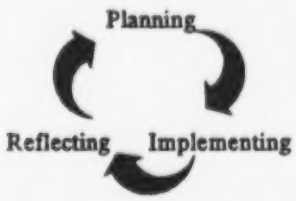
Organizations are being challenged to produce increasingly fast and fundamental changes to how they do business. The success of firms in executing high speed, organization-wide transformation is dependent upon their capacities for organizational learning. As a result, widespread interest exists in the creation of reliable and valid methods to assess organizational learning capabilities. While numerous diagnostic and assessment methods are being proposed, they primarily examine organizational characteristics and systems identified in the literature as associated with firms recognized as "learning organizations" (Gephart, Marsick, Van Buren, Spiro, and Lucadama, 1995). No direct and comprehensive assessment of organizational learning processes is currently known to exist. Such an assessment method should be able to make comparative judgments regarding the speed and effectiveness of organizational learning among different organizations, within the same organization over time, and among subunits of the same organization. This paper proposes such a framework and describes a pilot study of its application.

Strategic Learning Model

The assessment process presented here is based upon the Strategic Learning Model proposed by Redding and Catalanello (1994). Derived from best practice field study of leading change-oriented companies, this model offers an alternative to traditional strategic planning models. It proposes that the dynamic, iterative process through which organizations plan and execute strategic change can best be represented as follows:

1. Organizations formulate and execute strategic change through a series of organizational learning cycles, with each cycle comprised of three phases: planning, implementing, and reflecting. (See Figure 1.)

Figure 1. The Strategic Learning Cycle



© *copyright, John C. Redding, 1996*

2. Strategic learning cycles occur at three structural levels in the organization:

   Organization/business Unit Level
   Team Level
   Individual Level

3. Each strategic learning cycle can be assessed across three dimensions:

   Speed – The length of time it takes to complete an iteration of the learning cycle
   Depth – The degree to which single-loop and double-loop learning occurs during each iteration of the learning cycle. In contrast to single-loop learning, double-loop learning represents a deeper form of learning that is an increasing necessity for organizational survival amid chaotic, turbulent business conditions. According to Argyris (1992, p. 9), "Single-loop learning is appropriate for the routine, repetitive issues – it helps get the everyday job done. Double-loop learning is more relevant for the complex, non-programmable issues – it assures that there will be another day in the future of the organization."
   Breadth – The degree to which learning is transferred to other parts of the larger organizational system or to other strategic challenges facing the firm

According to the strategic learning model, the capacity of an organization to produce fast and fundamental strategic change is based upon the speed, depth, and breadth of strategic learning at the organization/business unit, team, and individual levels.

Strategic Learning Assessment Process

The Strategic Learning Model has the potential to provide a practical framework for assessing organizational learning processes as they contribute to the creation of strategic change. The following outlines the steps an assessment team might take to conduct such an assessment and illustrates how the assessment might be applied in the case of an office equipment manufacturing firm.

Step 1: Identify a Strategic Challenge. The assessment team begins by selecting a major strategic challenge that the organization has identified as critical to its future success. The actions taken to address this challenge will serve as the focus for the assessment at all three levels: the organization/business unit level, the team level, and the individual level. In the case of the equipment manufacturer, an assessment team might select the firm's planned North American consolidation as a strategic challenge to be assessed. During a recent strategic planning session, the company determined that competitive pressures are requiring that it consolidate its U.S. and Canadian operations into a single North American unit, while still meeting the unique needs of both markets.

Step 2: Track Actions Taken to Plan and Execute Change. Next, the assessment team tracks the actions taken in the organization to address the strategic challenge. These actions are followed at three structural levels: the organization/business unit level, the team level, and the individual level. At all three levels, information is specifically gathered to identify and evaluate iterations of the learning cycle as plans are developed, implemented, reflected upon, and new plans created as a result.

Step 3: Identify Learning Cycles. The actions taken at organization/business unit, team, and individual levels are examined to identify iterations of the learning cycle.

Organization/Business Unit Level. In many cases, learning cycles at this level occur over a multiyear period. For example, the equipment manufacturer developed a plan to consolidate its North American operations over a six month period. It then implemented the actions during the subsequent 12 month period. A year later, the organization reflected on the effectiveness of the North American consolidation, and, as a result, made changes to its plans. At this point, the first

learning cycle iteration was completed. These revised strategies were then implemented over the next year, initiating a second iteration.

Team Level. The planning and implementing of strategic change is often executed by teams acting on behalf of the whole organization. For example, to develop the initial North American consolidation plans, the equipment manufacturer established six teams: a marketing team, a financial team, a distribution team, an operations team, an information systems team, and a human resource team. Each team was charged with developing a plan of action to consolidate U.S. and Canadian operations in its assigned area and coordinating these plans with other teams. The actions taken by each team can be understood as a series of learning cycles. For example, the marketing team met for two days and identified the key issues that needed to be addressed to produce a successful consolidation. It left the meeting with a plan to investigate these issues over the next month. The team met again a month later, in which it shared the information that had been gathered and reflected upon what had been learned. At this point, a first iteration was completed. During the second meeting, the team identified a general strategy to consolidate U.S. and Canadian marketing efforts. Over the next month, these plans were presented to each of the other teams and to management. The team then met again and revised its plans based upon the feedback it received. At this point, a second iteration was completed. The team then developed a specific implementation plan to execute the strategy, thus starting a third iteration.

Individual Level. The actual implementation of strategic change often requires many members of the organization to make basic changes to their daily job responsibilities. The actions taken by individuals to execute the strategic change can similarly be understood as occurring through a series of learning cycles. For example, as part of the plan to consolidate North American marketing efforts, the vice president of marketing, U.S., now became the vice president of marketing, North America. After assuming this broader set of responsibilities, the individual was quickly surprised by the differences between advertising expenditures in the U.S. vs. Canada. In the U.S., advertising dollars were entirely devoted to a series of large national advertising campaigns. In Canada, the organization relied almost entirely upon direct mail marketing. The vice president decided to phase out Canadian direct mail over a six month period. Two months later, the vice president reviewed Canadian bookings and noticed that they had begun to decrease. Several managers felt that the decrease was a result of the phasing out of direct mail marketing. In response, the vice president decided to take a trip to Canada and visit a range of different customers. In response, the vice president decided to reinstate direct mail as the primary advertising method in Canada. The actions taken by the vice president can be understood as a series of learning cycles in which the vice president developed a plan of action, executed the plan, and reflected to determine what had been learned as a result of the actions.

Step 4: Assess the Speed, Breadth, and Depth of Strategic Learning. At all three levels, the assessment team then attempts to reach consensus on the speed, depth, and breadth of strategic learning.

Speed. The assessment team identifies dates for each of the following:

1. Start date and end date of planning phase for cycle #1 (P^1)
2. Start date and end date of implementation phase for cycle #1 (I^1)
3. Start date and end date of reflection phase for cycle #1 (R^1)

The assessment team then calculates the overall cycle time for the first learning cycle, as well as cycle times for each phase of the first cycle. This process is repeated for each learning cycle iteration that is being assessed.

Applying this process to the equipment manufacturer attempting to consolidate all North American operations, if sufficient information is available, the assessment team should be able to calculate the overall learning speed at the organization/business unit level as the firm executed the consolidation. Similarly, the team should also be able to calculate the learning speed at the team level for each of the six teams that developed the consolidation plans. Moreover, it should be able to calculate the learning speed of key individuals charged with implementing the plan, such as the vice-president of marketing.

Depth. For each iteration, the assessment team then compares the original plan of action completed during the first planning phase (P^1) with the plan of action completed during the second planning phase (P^2). The team then rates each iteration as follows:

No learning occurred – P^1 and P^2 are the same.
Single-loop learning occurred – P^1 and P^2 are different, and the two plans are based upon the same set of framing assumptions.
Double-loop learning occurred – P^1 and P^2 are different, and the two plans are based upon a different set of framing assumptions.

For an accurate assessment, the assessment team needs to reach consensus regarding the existence or nonexistence of framing assumptions. Framing assumptions are factors or conditions that need to be present for the plan to make sense. In most cases, these assumptions are not explicitly stated in the plan or by the planners. They can be inferred only by looking at the actions contained in the plan. For example, with the North American consolidation example, if the first plan was based on the assumption that Canadian customers expect the same levels of service as do U.S. customers and the second plan is based on the assumption that Canadian customers have different service expectations than do U.S. customers, one can conclude that double-loop learning occurred as a result of the first iteration.

Breadth. For each iteration, the assessment team then determines whether the learning that resulted from the learning cycle is transferred and used elsewhere in the organization. If so, it should be possible to determine whether plans developed in other parts of the organization have been modified as a result of the learning that has occurred as the organization, teams, and individuals take action to meet this strategic challenge. For example, let us say that one of the North American consolidation teams, the marketing team, discovers that customer expectations in Canada are significantly different than customer expectations in the U.S. If this learning is transferred to the other five teams, and the other five teams modify their plans as a result, one can conclude that a successful transfer of learning has occurred.

Pilot Study

A pilot study of the Strategic Learning Assessment Process is being conducted to answer several questions:

1. While the Strategic Learning Assessment Process appears theoretically possible, to what degree can it be practically applied?
2. Are certain structural levels – organization, team, or individual – assessed more or less effectively than others?
3. Are certain dimensions of learning – speed, depth, or breadth – assessed more or less effectively than others?

Three public elementary schools in Illinois were selected as sites for the study. Three years previously, each of the three schools had begun an intensive process of organizational transformation designed to make substantial improvements in the quality of teaching and learning in the school.

The study is being conducted in three phases. Phase 1 assesses the organization/business unit level, phase 2 assesses the team level, and phase 3 assesses the individual level. To date, only phase one has been completed.

The following is a description of the results of phase one, the organization/business unit level assessment. The assessment team consists of three individuals who were uninvolved in the change efforts in the three schools and who were knowledgeable of the Strategic Learning Model.

Step 1: Identify a Strategic Challenge. Three years ago, each of the schools had identified a small number of strategic challenges facing the school. For each school, the assessment team

selected one strategic challenge to be assessed. In school #1, the challenge was to improve student achievement in reading. In school #2, the challenge was to improve student behavior in the classroom. In school #3, the challenge was to improve student achievement in mathematics.

Step 2: Track Actions Taken to Plan and Execute Change. The assessment team reviewed records that documented the change process that has unfolded in each school over the three year period. They also interviewed individuals involved in the change initiatives in each school, using an interview protocol designed to provide information that could be used to identify learning cycles and assess the speed, depth, and breadth of learning.

Step 3: Identify Learning Cycles. The assessment team identified that one full learning cycle had been completed in each school. Two of the schools were in the implementation phase of the second iteration. One school had completed the planning phase of the second iteration.

Step 4: Assess the Speed, Breadth, and Depth of Strategic Learning. The team made the following assessments of the speed, depth, and breadth of strategic learning at the organization/business unit level:

Speed – The assessment team calculated the number of months it took each team to complete each phase of the first iteration and a total time for the first iteration as displayed in Table 1.

Table 1. Assessment of Speed of Learning at Organization Level in Three Elementary Schools

Organization	*Planning Phase*	*Implementation Phase*	*Reflection Phase*	*Learning Cycle Total*
School #1	5 months	12 months	10 months	27 months
School #2	4 months	8 months	1 month	13 months
School #3	12 months	9 months	1 month	22 months

The speed of learning varied from 13 months to 27 months among the three schools. There was also considerable variation among the time devoted to each phase of the learning cycle. For example, school #3 spent 12 months developing its original plan, compared to four months for school #2. School #1 spent 10 months in the reflection phase, compared to one month for the other two schools.

Depth. The assessment team assessed the depth of learning as displayed in Table 2.

Table 2. Assessment of Depth of Learning at Organization Level in Three Elementary Schools

Organization	*No Learning*	*Single-loop Learning*	*Double-loop learning*
School #1			X
School #2		X	
School #3		X	

In all three schools, the plans developed during the second iteration were different from the plans developed during the first iteration. In two of the three cases, the assessment team concluded that single-loop learning occurred. School #2 discovered that increased follow-up and accountability were needed to ensure that teachers actually made desired changes in the classroom. School #3 discovered that increased care had to be taken to select a consultant who could provide effective staff development and teacher coaching. Both of these learnings resulted in alterations of the plans without changing the framing assumptions upon which the plans were based. Only in school #1 did the assessment team conclude that double-loop learning occurred. In this case, the

school did not see desired improvement in reading scores as a result of the initial plan. The new plan placed an increased emphasis on reading comprehension. The assessment team concluded that this change reflected a shift in the framing assumptions regarding what is necessary to improve reading achievement in the school.

Breadth. The assessment team assessed the breadth of learning as displayed in Table 3.

Table 3. Assessment of Breadth of Learning at Organization Level in Three Elementary Schools

Organization	*Evidence of Learning Transfer?*
School #1	Undetermined
School #2	No
School #3	Yes

For school #1, the assessment team was unable to determine if, in fact, learning transfer had occurred. Other schools aware of the school's experience have begun to put an increased emphasis on reading comprehension. However, it was not clear to the assessment team whether these efforts directly resulted from school #1's learning. For school #2, there was no evidence that the team's learning has been applied elsewhere. For school #3, there was clear evidence that, both in school #3 and in other schools, its key learning had been successfully transferred.

Conclusions and Discussion

The first phase of the pilot study provided evidence that the Strategic Learning Assessment Process can be successfully applied at the organization/business unit level. For all three schools, the assessment team was able to reach consensus assessments of the speed and depth of learning. However, the assessment team experienced two difficulties in assessing the breadth of learning:

1. In most cases, numerous changes to the plans were made from one iteration to the next. In order to assess the breadth of learning, the assessment team needed to select one key learning that would be examined to determine if learning transfer occurred. There appeared to be some arbitrariness in the selection process.
2. There was limited information regarding the process through which transfer actually occurred. As a result, the team needed to speculate whether an actual transfer of learning did or did not occur. In one case, the team was unable to make the determination.

The pilot study suggests two improvements to the process:

1. Obtain detailed information as the learning cycle is occurring, such as learning logs, audiotape recordings of meetings, and observer notes.
2. Create a more detailed interview protocol to assist the assessment team in gathering information that will reduce the arbitrariness of some parts of the assessment.

In addition to completing phase two and phase three of the pilot study (which focus on the team learning and individual learning levels), future research is planned to apply the organization-wide assessment to a range of organizations of different size representing different industry groups and facing a variety of strategic challenges. Research is also planned to examine the interrater reliability of the assessment process. A parallel research process is currently investigating how the Strategic Learning Assessment Process can assist in the diagnosis of intervention opportunities and the selection of intervention approaches for team learning facilitators (Jeris, May, and Redding, 1996; Redding 1996).

References

Argyris, C. *Organizational Learning.* Cambridge, Mass.: Blackwell, 1992.

Jeris, L., May, S., & Redding, J. "Team Learning: Processes, Interventions and Assessments." Paper presented at the Annual Conference of the Academy of Human Resource Development, Minneapolis, Minnesota, March, 1996.

Gephart, M.A. Marsick, V.J., VanBuren, M.E., Spiro, M., & Lucadamo, L.I. *Learning Organization Assessment Framework.* Washington, D.C.: American Society for Training and Development, 1995.

Redding, J. *The Team Learning Consultant: Wrestling the Alligator of Change.* (Applied Research Report #2) Naperville, IL: Institute for Strategic Learning, 1996.

Redding, J., & Catalanello, R. *Strategic Readiness: The Making of the Learning Organization.* San Francisco: Jossey-Bass, 1994.

The Realities of Communicating and Reporting Practices and Outcomes for Internal and External Evaluators

Rosalie T. Torres
Torres Consulting Group

Hallie Preskill
University of New Mexico

Mary E. Piontek
RMC Research Corp.

This study investigated internal and external evaluators' practices and concerns about communicating and reporting evaluation findings. Approximately three-quarters (70%) of a random sample of American Evaluation Association members responded to a survey on this topic. Most evaluators adhered to traditional reporting formats. They were only moderately satisfied with their communicating and reporting efforts. Insufficient time and political/organizational complexity impeded success. Effective practice typically entailed high stakeholder involvement. Internal evaluation was found to be not only equally as prevalent as external evaluation, but different in important ways.

No aspect of evaluation is more fundamental than its use, in particular how we communicate about evaluation activities and report evaluation findings. Over the past 15 years investigations of the circumstances under which evaluations of projects, programs, and organizations are used by clients and stakeholders have yielded frameworks and models for explaining and facilitating evaluation use in a variety of settings (e.g., Alkin, 1985; Braskamp, 1982; Braskamp & Brown, 1980; Patton, 1986; Preskill, 1994, 1991; Torres, 1990; 1991). Much of this work has focused on how evaluations should be conducted to improve use, e.g., by attending to the specific context of the evaluation, by involving clients and stakeholders in the evaluation design, by reporting results in a timely manner. While specific factors about evaluation communicating and reporting have been addressed in the literature, there has been no systematic, comprehensive treatment of this topic. Moreover, what has been addressed has not been clearly grounded in problems and issues currently faced by evaluators practicing in a variety of settings.

In these settings evaluators face circumstances which continue to speak to the need for yet better understanding of how evaluation practice can be made more efficient and effective. For example, evaluators are experiencing increased : (a) demands for accountability and effectiveness in education and business domains; (b) change and complexity within organizations; and (c) acceptance and use of non-traditional evaluation methods (i.e., case study, ethnographic, qualitative designs) which require different, innovative reporting styles.

This study investigates internal and external evaluators' current practices and concerns about communicating and reporting evaluation findings. It addressed the following specific questions through a survey of members of the American Evaluation Association (AEA):

1. What methods and styles do evaluators use and with what frequency?
2. How satisfied are evaluators with their communicating and reporting efforts?
3. What factors most impede success in communicating and reporting?
4. Are specific practices for, experiences with, and perceptions about communicating and reporting different for internal and external evaluators?
5. What kinds of communicating and reporting activities have been most successful?
6. What kinds of practical experiences have been most instructive to evaluators about communicating and reporting?

© copyright, Torres, Preskill, & Piontek, 1996

BEST COPY AVAILABLE

Methodology

Respondents The response rate for the survey was 70 percent (240), calculated on the basis of 343 possible randomly selected respondents. Of these 240, 16 (7%) identified themselves as never having been involved in conducting evaluations of programs, projects, or organizations. As requested, they returned the survey without completing the remainder of it. Of the 224 respondents who had conducted evaluation in the past, 33 (15%) identified themselves as not currently involved in conducting evaluations. These respondents then only answered portions of the survey focusing on prior evaluation activities.

The remaining 191 respondents currently working as evaluators were approximately equal split between internal evaluators (51%), and external evaluators (45%) employed in a variety of organizations. Overall, one-third are employed in higher education, equally split between internal and external evaluators. Next in frequency are external evaluators working for consulting firms (including self-employment, 15%), and internal evaluators in non-profit organizations (12%). Only seven percent work for school systems, and almost all of these individuals are internal evaluators. The fewest numbers of respondents are employed in business and industry, local government, and health care organizations. Again, the majority of these individuals are internal evaluators.

The respondents' total years in evaluation practice represents an approximately normal distribution. That is, 27 percent have been conducting evaluations from less than one year to six years, 51 percent have been conducting evaluations from seven to fifteen years, and 22 percent have been conducting evaluations for 16 years or more.

Instrumentation and Data Analysis In addition to demographic items reported above, the survey included (a) Likert-scale items on evaluation practices, formats, audiences, and satisfaction with communicating and reporting efforts; (b) a checklist on factors which impede successful communicating and reporting; and (c) a critical incident question about respondents' most instructive/successful experiences with communicating and reporting. Of the total number of respondents, 61 percent (146) provided responses to the critical incident item. The illustrative cases they wrote varied in detail and length from 5 to 88 lines of text.

Descriptive statistics were calculated for 59 quantitative survey items using SPSS (Statistical Package for the Social Sciences). In addition, Mann-Whitney U tests were run to compare the responses of internal and external evaluators on the frequency of (a) their use of various communicating and reporting practices and (b) their reporting to various audiences. The illustrative cases were analyzed using The Ethnograph Software (Qualis Research Associates, 1987) for text-based data. In all, 110 codes in eight categories were generated for the analysis of these data. The categories are (a) type of program evaluated, (b) approach/focus/questions of the evaluation, (c) role of the evaluator, (c) communicating/reporting formats used, (d) audiences addressed, (e) successful formats/processes used, (f) factors impeding to success, and (g) outcomes of successful reporting. These categories and codes were modified in three iterative cycles to maximize their fit with the data.

Summary of Findings

The following points summarize our major findings. First, to communicate and report the findings of their work, practicing evaluators most frequently use traditional methods (see Table 1). That is, they: (a) write technical/final reports, (b) develop written evaluation plans prior to implementation, (c) write executive summaries, (d) specifically identify evaluation audiences, (e) make formal verbal presentations to clients and/or staff, (f) develop reporting plans.

Second, external evaluators tend to engage in these practices more frequently than internal evaluators. Also, external evaluators more frequently (a) conduct planned personal discussions in person or by telephone, (b) make conference presentations, and (c) submit reports of evaluations for publication. (See Table 1.)

Table 1
MEAN FREQUENCY OF EVALUATORS' COMMUNICATING/REPORTING PRACTICES IN THEIR CURRENT POSITION BY TYPE OF EVALUATOR

Communicating/Reporting Practice	Overall		Type of Evaluator: Internal		External	
	Mean[a]	N	Mean	N	Mean	N
(Write) technical/final report	3.7	181	3.5	96	3.8**	85
Develop written evaluation plan prior to implementation	3.5	183	3.4	98	3.6*	85
(Write) executive summary	3.5	182	3.3	97	3.6**	85
Specifically identify evaluation audiences (either as part of the evaluation plan or separately)	3.2	182	3.2	98	3.3	84
(Make) formal verbal presentation to clients/staff	3.1	183	2.9	98	3.2*	85
Develop reporting plan (either as part of the evaluation plan or separately)	3.0	183	2.8	98	3.3**	85
(Conduct) working session (e.g. with staff for action planning /interpretation of findings)	2.9	183	2.8	98	2.9	85
(Hold) impromptu personal discussions (in person or by telephone)	2.8	183	2.6	98	2.9	85
(Write) short memo(s)/progress reports	2.6	183	2.6	98	2.7	85
(Hold) planned personal discussions (in person or by telephone)	2.5	183	2.3	98	2.8*	85
(Write) interim reports	2.5	182	2.3	97	2.6	85
(Make) conference presentation	1.9	183	1.7	98	2.2**	85
(Publish) journal article	1.3	182	1.1	97	1.6**	85
(Write for) newsletters internal to the organization/program	1.1	181	1.2*	97	.8	84
(Participate in) public meeting	1.1	180	1.0	95	1.1	85
(Provide) news release	1.0	182	.8	98	1.0	84
(Write for) newsletters external to the organization/program	.8	181	.9	97	.7	84
(Use) brochure	.6	182	.7	97	.6	85
(Use) photographs	.5	182	.5	97	.5	85
(Make) video tape	.4	181	.5	96	.3	85
(Hold) press conference	.4	182	.4	97	.4	85
(Make) television/radio appearance	.4	179	.3	95	.6*	84
(Use) skit/psychodrama	.06	182	.06	97	.04	85

Note. Response scale ranged from 0 to 4 where 0 = never and 4 = more than 75% of the time.
[a] Levels of significant differences between mean frequencies for internal and external evaluators are indicated as follows: * - .05 and ** - .01.

BEST COPY AVAILABLE

Third, for both groups, little, if any, time is spent communicating findings through less traditional means: public meetings, external newsletters, brochures, photographs, videotapes, press conferences, television/radio appearances, and skits or psychodramas.

Fourth, the single greatest impediment to successful communicating and reporting is insufficient time to devote to this aspect of the evaluation endeavor (see Table 2). Political and organizational complexity poses another major impeding factor. Specifically, evaluators found the following aspects of conducting their work within organizations difficult: (a) lack of clarity among stakeholders about communicating/reporting needs, (b) unresponsiveness to communicating/reporting efforts, (c) client/audience turnover, (d) politically charged situations, (e) resistance to negative findings, (f) characteristics of particular individuals, and (g) misinterpretation of findings.

Fifth, evaluators are only moderately satisfied with their communicating and reporting efforts and outcomes. But, more experienced evaluators and external evaluators are more satisfied than other evaluators are. Yet, the longer evaluators are in their current positions, the less satisfied they are with their communicating and reporting (see Table 3).

Sixth, evaluators, however, did identify successful elements for each aspect of communicating and reporting--format, content, and processes. They described the importance of using various *formats* including short reports and summaries tailored to audience needs. The *content* of successful communicating and reporting efforts included clear language; graphs and charts; positive as well as negative findings; qualitative, contextual data as well as quantitative data; and specific recommendations. Ongoing, collaborative communication *processes* were the most successful. Respondents stressed stakeholders' involvement in the conduct of the evaluation as a whole, but especially in interpreting findings.

Discussion

These findings about actual communicating and reporting practices lead us to five major conclusions. First, it confirms much of what we have known to be good practice in evaluation--attending to the specific context of an evaluation, involving clients and stakeholders in the evaluation design, and reporting results in a timely manner to a variety of audiences using clear language. Success comes through communication and collaboration throughout the evaluation process, and from the presentation of information in such a way that it is easily assimilated. Communicating and reporting is part and parcel of the entire evaluation endeavor--not something undertaken at the end of the evaluation process. Second, we found that while some evaluators successfully implement these strategies, many others remain thwarted in their efforts by time, organizational, and political constraints. And, not surprisingly, evaluators are only moderately satisfied with their efforts. This leads to our third conclusion: Satisfaction with, and likely skill at, communicating and reporting comes with experience.

Fourth, we have found the practice of internal evaluation to be not only equally as prevalent as external evaluation, but also different in important ways. Comparing the two makes even clearer the challenges posed in working with complex, constantly changing, politically charged organizations. The meaning and implications of these conclusions are discussed further in terms of five major issues below.

Collaboration throughout the Evaluation The results of this study--particularly those from the illustrative cases--make clear that issues related to communicating and reporting evaluation findings are of concern throughout the evaluation process, not just at the time that end-products of the evaluation such as final reports are typically expected. Early, ongoing communication and collaboration were identified by evaluators in two ways: first, as strategies which would have improved, if not prevented, frustrating experiences they've had in communicating and reporting; and second, as what caused things to go right when they did. This finding echoes Patton's long-held position (1978, 1986) that evaluation use must be planned for at the beginning of, and attended to, throughout the evaluation.

BEST COPY AVAILABLE

Table 2
FREQUENCY WITH WHICH EVALUATORS INDICATED FACTORS IMPEDING SUCCESS IN COMMUNICATING AND REPORTING EVALUATION FINDINGS

Factors Impeding Success in Communicating and Reporting Evaluation Findings	Percent	N
Insufficient time available to devote to communicating/reporting	53%	116
Clients/audiences unclear about their communicating/reporting needs	46%	100
Client/audience unresponsiveness to communicating/reporting efforts	39%	87
Changes in audiences, organizational/program leadership	38%	86
Insufficient time for analysis/interpretation of findings	36%	81
Insufficient resources available to cover personnel cost for communicating/reporting	33%	73
Inaccurate/incomplete analysis of political context	27%	59
Too many different formats necessary for communicating/reporting to different audiences	26%	58
Misuse/misinterpretation of formal communications/reports by clients/audiences	24%	53
Lack of planning for effective communicating/reporting	23%	52
Misuse/misinterpretation of informal communications/reports by clients/audiences	18%	41
Difficulty in balancing positive and negative findings	17%	38
Language/style used in communicating/reporting not clearly understood by audiences	15%	34
Difficulty in integrating findings across multiple sites	13%	29
Inadequate/insufficient training in communicating and reporting	11%	25
Insufficient resources available to cover materials costs for communicating/reporting	10%	23
Insufficient resources available to cover technology costs for communicating/reporting	10%	22

Table 3
EVALUATORS' MEAN LEVEL OF SATISFACTION WITH COMMUNICATING AND REPORTING EVALUATIONS BY YEARS CONDUCTING EVALUATIONS AND YEARS IN PRESENT POSITION

	Conducting Evaluations		In Present Position	
Years	Mean	N	Mean	N
Less than 1 year	2.5	2	4.2	9
1 - 3 years	3.0	19	3.4	48
4 - 6 years	3.3	36	3.4	41
7 - 10 years	3.3	56	3.4	37
11 - 15 years	3.5	53	3.4	29
16 - 20 years	3.8	26	3.6	13
20 + years	3.8	25	3.2	9
TOTALS	3.3	217	3.4	186

BEST COPY AVAILABLE

Communication and collaboration takes time--the single factor most plainly implicated in this study. For these evaluators time well spent would find them (a) communicating with participants and stakeholders of varied levels about evaluation purposes and processes; and (b) involving these individuals in the design, data collection, analysis, interpretation, and follow-up phases of the evaluation. They found that doing so informed and educated clients/audiences about evaluation, sensitized evaluators to individual perspectives and organizational contexts, clarified mutual expectations, and increased client/audience capacity for using the evaluation findings--whether it be to shape their understanding of the program or to make a specific decision. In short, their responses confirmed that involving clients/audiences in a well-conceived evaluation process increases the credibility of the effort, enhances audiences' understanding of both evaluation and the program, and increases the possibility that reflection and action will follow.

Communication and collaboration alone, however, did not account for the evaluators' successes. Evaluators had provided information in a form that could be quickly and easily assimilated--using verbal presentations, discussion, and interaction at meetings and in informal exchanges; short, cogent reports; clear, jargon-free, familiar language; and graphics and other formats which relate and integrate information visually. The content of these communications included quantitative and qualitative information, descriptions of context, positive and negative findings, and specific recommendations.

Social Science Research Orientation The single major impediment to successful communicating and reporting cited by evaluators is time. Of particular concern is the time it takes for the analysis and interpretation of findings. Moreover, analysis and interpretation of findings is a consuming aspect of developing final reports--which is the most frequently practiced means of communicating and reporting findings. On average, respondents in our study indicated they write final/technical reports between 75 and 100 percent of the time they do evaluations. Indeed, typical expectations dictate that evaluators generate lengthy, comprehensive final reports. In this regard, current evaluation practice is following social science traditions and basic research training most evaluators received and continue to provide students of evaluation. This tradition is generally reflected in many texts on evaluation (see e.g., Popham, 1993; Posavac & Carey, 1992; Worthen & Sanders, 1987) as well as The Program Evaluation Standards (Joint Committee, 1994). Scriven (1993) explicitly links evaluation reports which look like social science research reports with attempting to show that program evaluation is applied social science. He finds this format "almost useless for most clients and audiences" (p. 77). Alkin (1990, p. 160) worries "about the extent to which...[evaluators] become so involved...with the end result of producing a report for peer review, that [they] fail to serve the evaluative needs of the situation."

Likewise, our findings strongly suggest that evaluators may be spending too much time generating such reports in lieu of other reporting methods. It is unclear that such reports provide the most cost effective means for communicating evaluation findings in ways most useful to stakeholding audiences.

Given the heavy reliance on final/technical reports and evaluators' dissatisfaction with their communicating and reporting efforts and outcomes in general, it is worthwhile to consider the potential of using non-traditional communicating and reporting approaches for different types of evaluation activities. Important questions for evaluators to consider are: Should formative evaluation studies result in long, summative reports? Would other methods--photographs, debriefing sessions and newsletters, for instance--lead to greater instrumental and conceptual use? As one respondent commented, "Frankly, we have more creative ideas about reporting/communicating than we have time to implement. Time and resources are constraining." Our greatest challenge may be to give up the familiar for the new in our practice.

Political and Organizational Complexity Let's turn now to the other major obstacle to successful communicating and reporting--political/organizational complexity. The majority of impediments evaluators frequently cite relate to working with the individuals and organizations of the programs being evaluated. The issues facing evaluators include lack of clarity about needs for communicating and reporting, unresponsiveness, organizational and personnel changes, and inaccurate or incomplete analysis of the political context. Evaluators specifically described

BEST COPY AVAILABLE

politically charged situations dealing with hidden agendas, turf protection, and gatekeeping. They detailed instances where resistance to negative findings hindered their efforts, and in some cases misinterpretation of results prevailed. Finally, our respondents sometimes implicated specific individuals who lacked training, competence, and/or political savvy.

Yet, some evaluators have been able to manage these challenges by taking a more collaborative approach, useful in mediating these complexities. Collaboration in evaluation is powerful: it can enhance use by different audiences, it values individuals' experiences and opinions, it can contribute to a sense of ownership, it can produce better understanding and depiction of the context, it can lead to more useful recommendations, it can educate clients/audiences about the program and evaluation, and it can help identify and resolve conflicts before the end of an evaluation. It embraces different perspectives and lets many voices be heard.

Internal and External Evaluation Approximately one-half of the 191 evaluators responding to our survey who currently practice evaluation are internal evaluators. While we do not know how much internal evaluation has been practiced in the past, these results suggest that it is now at least equally as prevalent as external evaluation. This equity is important to consider since most evaluation models used in the training of evaluators were developed from the perspective of external evaluators (see e.g. Guba & Lincoln, 1990, Scriven, 1973; Stufflebeam, 1983; Stake, 1983).

While the existing literature tells us "something about internal evaluation...there remains a great deal to be described and understood" (Mathison, 1991a, p. 164). This study contributes to that understanding by documenting clear differences between the practice of internal and external evaluation. Internal evaluators less frequently engage in many practices commonly expected to facilitate use--namely, developing evaluation and reporting plans, writing executive summaries, making formal verbal presentations, and holding planned personal discussions (see Shadish, Cook, & Leviton, 1991; Patton, 1986). Relatedly, internal evaluators in our study cited lack of planning for communicating and reporting significantly more frequently than did external evaluators as an impediment to success.

These circumstances are likely, in part, to be caused by the organizational and role complexity internal evaluators experience (Mathison, 1991b). Increased demands on internal evaluators to respond to organizational needs and circumstances mitigates time and energy they might otherwise have for more effective communicating and reporting practices. Moreover, internal evaluators are significantly less satisfied with their communicating and reporting efforts than are external evaluators. The issue is not so much that different practices are more effective for one type of evaluation than the other; but rather, internal and external evaluators face different challenges in carrying out effective communicating and reporting strategies.

Satisfaction Comes with Experience Given what we have learned from our research on communicating and reporting, it makes some sense that while satisfaction is generally low, it does increase with experience. Only after at least four years of practice did our respondents' overall satisfaction shift, albeit slightly, to the positive. Increased skills and confidence likely contribute to greater satisfaction. With more experience evaluators better understand various organizations and audiences. In spite of this learning, however--time, resources, and political and organizational complexity--still pose challenges to the success of communicating and reporting efforts. Here again we find affirmation of what evaluators have now long discussed among themselves: Evaluation is a complex, interdependent, demanding, and challenging endeavor. Yet, rather than be discouraged, we find ourselves emboldened--striving to seize the opportunity by comprehending those predicaments we face and our role in them.

Based on our findings about what works and doesn't work, continued discussion and study among ourselves of what it means to practice both internal and external evaluation, how we can train future evaluators, how we can make evaluation information easier to assimilate, and how we can enrich evaluation practice and outcomes for both clients and evaluators is more vital than ever (see Torres, Preskill, & Piontek, in press).

References

Alkin, M.C. (1985). A guidebook for evaluation decision-makers. Newbury Park: Sage.

Alkin, M.C. (1990). Debates on evaluation. Newbury Park: Sage.

Braskamp, L.A. (1982). A definition of use. Studies in Educational Evaluation, 8, 169-174.

Braskamp, L.A., & Brown, R.D. (eds.) (1980). Utilization of evaluative information. New Directions for Program Evaluation, No. 5.

Cummings, O.W., & et.al. (1988). Business perspectives on internal/external evaluation. New Directions for Program Evaluation: Evaluation Utilization, 39, 59-73.

Guba, E.G., & Lincoln, Y.S. (1990). Fourth generation evaluation. Newbury Park: Sage.

House, E.R. (1986). Internal evaluation. Evaluation Practice, 7, 63-4.

Joint Committee on Standards for Educational Evaluation. (1994). The Program Evaluation Standards. Newbury Park: Sage.

Love, A.J. (1983). The organizational context and the development of internal evaluation. New Directions for Program Evaluation: Developing Effective Internal Evaluation.

Mathison, S. (1991a). What do we know about internal evaluation? Evaluation and Program Planning, 14(1), 159-165.

Mathison, S. (1991b). Role conflicts for internal evaluators. Evaluation and Program Planning, 14(1), 173-179.

Patton, M.Q. (1978). Utilization-focused evaluation. Beverly Hills: Sage.

Patton, M.Q. (1986). Utilization-focused evaluation (2nd ed.). Beverly Hills: Sage.

Popham, W.J. (1993). Educational evaluation. (3rd ed.). Needham Heights, MA: Allyn and Bacon.

Posavac, E.J. & Carey, R.G. (1992). Program evaluation: methods and case studies. (4th ed.). Englewood Cliffs: Prentice-Hall, Inc.

Preskill, H.S. (1991). The Cultural Lens: Bringing utilization into focus. In C.L. Larson & Preskill (Eds.), Organization in transition: opportunities and challenges for evaluation, New Directions for Program Evaluation, 49, 5-16.

Preskill, H. S. (1994). Evaluation's role in enhancing organizational learning: A model for practice. Evaluation and Program Planning, 17(4), 291-297.

Qualis Research Associates (1987). The Ethnograph Software. Corvallis, OR.

Scriven, M. (1993). Hard-won lessons in program evaluation. San Francisco: Jossey-Bass.

Shadish, W.R. Jr., Cook, T.D., & Leviton, L.C. (1991). Foundations of programs evaluations: Theories of practice. Newbury Park: Sage.

Stake, R.E. (1983). Program evaluation, particularly responsive evaluation. In G.F. Madaus, M.Scriven, & D.L. Stufflebeam (eds.), Evaluation models (287-310). Norwell, MA: Kluwer.

Stenzel, N. (1991). Evaluation in a state education agency. Evaluation and Program Planning, 14, 181-7.

Stufflebeam, D.L. (1983). The CIPP model for program evaluation. In G.F. Madaus, M. Scriven, & D.L. Stufflebeam (Ed.s), Evaluation models (117-142). Norwell, MA: Kluwer.

Torres, R. T. (1990). The role of context in the practice and use of internal evaluation activities: A case study from business and industry. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Torres, R.T. (1991). Improving the quality of internal evaluation: The consultant-mediator approach. Evaluation and Program Planning, 14(4), 189-198.

Torres, R. T., Preskill, H., & Piontek, M. E. (in press). Evaluation strategies for communicating and reporting: Enhancing learning in organizations. Newbury Park, CA: Sage.

Winberg, A. (1991). Maximizing the contributions of internal evaluation units. Evaluation and Program Planning, 14(1), 167-172.

Worthen, B.R. & Sanders, J.R. (1987). Educational evaluation: Alternative approaches and practical guidelines. White Plains, N.Y.: Longman Inc., 341-365.

BEST COPY AVAILABLE